Modeling Hydrophobic Solvation of Nonspherical Systems: Comparison of Use of Molecular Surface Area with Accessible Surface Area

ROBERT B. HERMANN

Department of Chemistry, Indiana University–Purdue University at Indianapolis, 402 North Blackford Street, Indianapolis, Indiana 46202

Received 9 January 1996; accepted 11 April 1996

ABSTRACT

It is shown that the molecular surface and the accessible surface lead to exactly the same results when calculating solvation free energies and transfer free energies, from methods using the surface tension as a parameter if the exact geometric curvature is used with the accessible surface. However, the use of the exact curvature is not necessarily the best approach chemically. Other modifications, including an approximate curvature improves the approach. Such modifications are difficult to include in methods in which the molecular surface rather than the accessible surface is used to calculate solvent effects. A modification of a Gaussian curvature term is necessary if dissociation is to be accounted for properly. The inclusion of a Gaussian curvature term, in addition to the usual mean curvature term, reconciles the difference in magnitude of the microscopic and macroscopic surface tension in the case of the accessible surface area.

Introduction

In the attempt to model the Gibbs free energy of hydrophobic solvation and of hydrophobic association, it is desirable to account for differences in molecular shape. In the usual attempts to account for hydrophobic or other solvation effects, two quantities of general interest are the accessible surface area[1] and the molecular surface area,[2] illustrated in Figure 1. The accessible surface follows naturally from the concept of a collision diameter, a concept firmly ingrained in the chemistry of molecular interactions. On the other hand, the molecular surface area for the treatment of solubilities has been regarded by some as more easily applied.[3–5] The use of a more informal

Journal of Computational Chemistry, Vol. 18, No. 1, 115–125 (1997)

CCC 0192-8651 / 97 / 010115-11

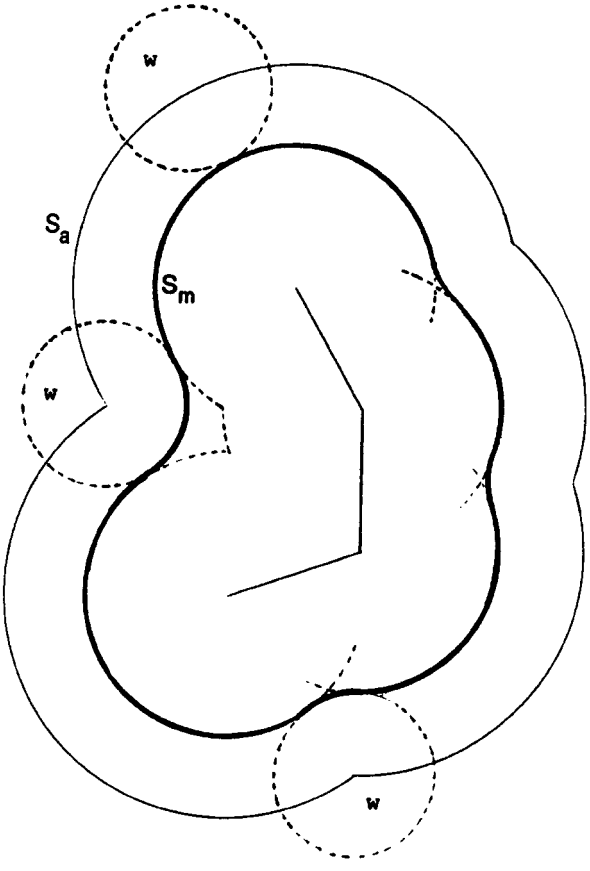

FIGURE 1. Inside heavy line: molecular surface; outside light line: accessible surface; w: water (solvent) molecules.

molecular surface goes back originally to Langmuir[6] and a van der Waals surface was later used by Valvani et al.[7] to treat solubilities.

Several investigators have demonstrated, through a correlation of solubility data with the molecular surface area, A_m, that the transfer free energy from hydrocarbon to water is directly proportional to the molecular surface area, and differences in surface curvature play no role.[3,4] The proportionality constant appears to be close to the macroscopic interfacial surface tension, γ_{int}, so that

$$\Delta G = A_m \gamma_{\text{int}} \tag{1}$$

By an appeal to scaled particle theory, Jackson and Sternberg[4] suggested that the free energy of cavity formation is independent of curvature, and therefore molecular shape, in the case of the molecular surface. On the other hand, in the case of the accessible surface, a curvature correction to the surface tension of a planar surface or interface, must be made. Sharp et al.[8] use a curvature correction to the surface tension

$$\gamma/\gamma_{\text{int}} = 1/(1 + a/R) \tag{2}$$

where $1/R$ is the local curvature at the molecular surface and a is the solvent radius. The form of this is consistent with the Tolman expression with $a = 2\delta$ for the ratio of the surface tension of water droplets to the macroscopic surface tension of water.[9]

Eq. (2) can be rewritten as

$$\gamma = \gamma_{\text{int}}(1 - a/(R + a)) \tag{3}$$

in terms of the collision diameter $r = R + a$ between a solute atom center and a solvent molecule.

In this article, the symbol, a, for the radius of a solvent will be used to be consistent with recent reports on this subject,[4,10] even though in most studies on scaled particle theory, a refers to the solvent diameter.[11–13]

The relationship between the two methods of determining the solution free energy and free energy of transfer are examined in what follows from two approaches, geometrical and physical. It is proven that, under certain conditions, the two methods are equivalent. The molecular area is geometrically related to the accessible area in a precise way involving a particular curvature correction. While the molecular surface allows for a good, simple linear correlation in many cases, it may not always be the best one to use. It is relatively easy to correlate hydrocarbon solubilities but difficult to generalize to more complex cases, such as those of biological interest. Hydrophobic interactions, aggregation, and binding to enzyme pockets are much more difficult to account for. There appear to be some physical difficulties which only an explicit consideration of curvature or some similar function can address.

Geometrical Considerations

To compare solvation free energies calculated by the two techniques, one using the molecular surface, A_m, and the other using the accessible surface, A_a, two equations are compared. Two possible expressions for ΔG_c, the free energy of cavity formation in the solvent, are

$$\Delta G_c = A_m \gamma_\infty \tag{4}$$

$$\Delta G_c = A_a \gamma_\infty g(\kappa) \tag{5}$$

where γ_∞ is the macroscopic surface tension and with similar expressions for the transfer free energy problem. The surfaces here are not necessarily spherical and $g(\kappa)$ is some function of the curvature. The curvature term, $g(\kappa)$, may be regarded as applying to the surface rather than to the surface tension. The basic problem then is to determine the form of the curvature term, $g(\kappa)$, which makes these two expressions equivalent. That is, a $g(\kappa)$ is needed such that

$$A_m = A_a g(\kappa) \tag{6}$$

The relationship between the molecular surface area and the accessible surface area may be examined in terms of the concept of parallel surfaces.

Parallel surfaces may be defined in the following way. Given a surface, S_m, defined parametrically by the vector function, $X(u, v)$, and a second surface S_a, defined parametrically by the vector function, $Y(u, v)$, then S_a is parallel to S_m if the distance of S_a from S_m as measured along a normal to S_m is a constant. If the distance of S_a from S_m as measured along a normal to S_m is a, then for any point, $X(u, v)$ on the surface S_m

$$X(u, v) = Y(u, v) + aN(u, v) \tag{7}$$

where $N(u, v)$ is a normal of unit length directed toward the interior of the surface (Fig. 2).

Taking the coordinates u and v in the surface as principal lines of curvature, it can be shown[14, 15] that at any point on the surface the vector product of the derivatives of $X(u, v)$ with respect to u and v are related to the vector product of the derivatives of $Y(u, v)$ by a formula involving the mean curvature and Gaussian curvature of S_a:

$$\frac{\partial X(u,v)}{\partial u} \times \frac{\partial X(u,v)}{\partial v} = \frac{\partial Y(u,v)}{\partial u} \times \frac{\partial Y(u,v)}{\partial v}\left(1 - 2a\kappa_m + a^2\kappa_g\right) \tag{8}$$

where the mean curvature may be defined as

$$\kappa_m = (1/r_u + 1/r_v)/2 \tag{9}$$

and the Gaussian curvature may be defined as

$$\kappa_g = (1/r_u)(1/r_v) \tag{10}$$

where r_u and r_v are the radii of curvature along the principal lines of curvature u and v, respectively. A principal radius r_u or r_v will be positive if its center of curvature is inside the surface and negative otherwise, because the direction of the surface normal was chosen to be toward the inside, as in Figure 2.

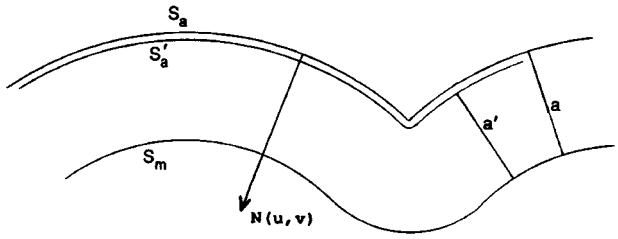

FIGURE 2. Accessible surface and molecular surface are parallel surfaces. Accessible surface, S_a, is the limit of the surface S'_a as $a' \to a$.

In the case of the accessible surface, there is a mathematical difficulty in that derivatives are not continuous where the spherical portions of the surface intersect, but by defining the surface as a limiting case, the above formulas apply. Referring to Figure 2, as the distance a' goes to a, the surface S'_a approaches the accessible surface S_a, and S'_a may be made as close to S_a as desired:

$$S_a = \lim_{a' \to a} S'_a \tag{11}$$

The above result may be used to relate the two surface areas, A_m and A_a. The surface area is the integral of the absolute value of the corresponding vector product over the entire surface, so that if A_a is the area of S_a

$$A_a = \int_{S_a} \left| \frac{\partial Y(u,v)}{\partial u} \times \frac{\partial Y(u,v)}{\partial v} \right| du\, dv \tag{12}$$

The average value of the quantity $1 - 2a\kappa_m + a^2\kappa_g$ in eq. (8) on the surface is

$$1 - 2a\overline{\kappa_m} + a^2\overline{\kappa_g} = \frac{\int \left(1 - 2a\kappa_m + a^2\kappa_g\right) \left| \frac{\partial Y(u,v)}{\partial u} \times \frac{\partial Y(u,v)}{\partial v} \right| du\, dv}{\int \left| \frac{\partial Y(u,v)}{\partial u} \times \frac{\partial Y(u,v)}{\partial v} \right| du\, dv} \tag{13}$$

Taking the absolute value of both sides of eq. (8), noting that the quantity $1 - 2a\kappa_m + a^2\kappa_g$ is always positive because r_u and r_v are always greater than

a or negative, and substituting the result into the integral in the numerator of eq. (13) gives

$$A_m = A_a\left(1 - 2a\overline{\kappa_m} + a^2\overline{\kappa_g}\right) \tag{14}$$

where $\overline{\kappa_m}$ and $\overline{\kappa_g}$ are the averages of the mean and Gaussian curvatures, respectively, averaged over the entire surface, and A_m is the area of S_m. So, $g(\kappa)$ is then

$$g(\kappa) = \left(1 - 2a\overline{\kappa_m} + a^2\overline{\kappa_g}\right) \tag{15}$$

and eq. (5) becomes

$$\Delta G_c = A_a\gamma_\infty\left(1 - 2a\overline{\kappa_m} + a^2\overline{\kappa_g}\right) \tag{16}$$

The Gaussian curvature, integrated over the whole surface is 4π for any surface topologically equivalent to a sphere, so that the above formula can be rewritten

$$\Delta G_c = A_a\gamma_\infty\left(1 - 2a\overline{\kappa_m}\right) + 4\pi a^2\gamma_\infty \tag{17}$$

for such surfaces. Finally, for spherical solutes of radius R, $\overline{\kappa_m} = 1/r$ so that

$$\Delta G_c = 4\pi r^2\gamma_\infty(1 - 2a/r + a^2/r^2) \tag{18}$$

where $r = (R + a)$.

The above results mean that the same value may be obtained for the free energy of cavity formation using either surface, provided the above particular curvature correction of eq. (15) is used in conjunction with the accessible surface. A general algorithm for accurately calculating the mean curvature of accessible surfaces of multicentered systems is given in Appendix A.

Microscopic Versus Macroscopic Interfacial Tension

If the curvature correction $g(\kappa)$ is regarded as applying to the surface tension and interfacial tension, rather than the accessible surface area, then the microscopic interfacial tension can be found from

$$\gamma = \gamma_{\text{int}}\left(1 - 2a\overline{\kappa_m} + a^2\overline{\kappa_g}\right) \tag{19}$$

where $\overline{\kappa_g} = 4\pi/A_a$. These results are directly applicable to the problem of the difference between the microscopic and macroscopic surface tension.[8,16]

For example, using 72–75 cal mol^{-1}Å^{-2} for the macroscopic interfacial tension γ_{int},[17] and curvatures from Table I, the microscopic value for the interfacial tension for *n*-butane is 32–33 cal mol^{-1}Å^{-2}. This may be compared with previous results by transfer free energy methods.[18–20] Thus, the inclusion of the Gaussian curvature term, in addition to the usual mean curvature term, accounts for the difference between the macroscopic and microscopic values. This is in line with the results of Jackson and Sternberg[4] who found that the macroscopic surface tension could be used with no curvature correction when applied to the molecular surface.

Physical Considerations

While the two surfaces are related geometrically as above, the curvature expression of eq. (16) may not be the best one physically. A curvature expression must give good results for ΔG_c, not only for simple hydrocarbon solutes, but also in systems with more complex or highly curved surfaces.

To investigate this point, several modifications were made. The results are given in Table I, for a united atom treatment, and compared to an experimentally derived result for ΔG_c:

$$\Delta G_c(\text{exp}) = \Delta G^* - e_i \tag{20}$$

where ΔG^* is the experimental solution free energy[21] (unitary) and e_i is calculated from OPLS potentials[22] and molecular dynamics as previously described.[23]

The simplest modification is to treat the coefficients of $\overline{\kappa_m}$ and $\overline{\kappa_g}$ in eq. (16) as adjustable parameters. The best values were found to be -2.67 and 1.33, as given in Table I.

Second, the use of the exact geometrical average mean curvature, $\overline{\kappa_m}$, may not be as nearly correct physically for nonspherical surfaces as it is for spherical surfaces, where it reduces to $1/(R + a)$ as in scaled particle theory.

Figure 3 gives an extreme example of a nonspherical surface. In this case, some regions of the surface are between one and two solvent diameters apart. The water molecules forming the surface at P_1 are at a different energy that at P_2, since the waters at P_2 are bound to more bulk water than at P_1. Using the exact $\overline{\kappa_m}$ in this idealized example, the surface at P_1 would incorrectly contribute the same surface energy as the surface at P_2, since the energy of the accessible surface would be unaffected by the distance to the molecular surface at P_3.

TABLE I.[a]
Cavity Energies and Associated Quantities.

Hydrocarbon	ΔG_c from exp.[c]	ΔG_c from exact[d] $\overline{\kappa_m}, \overline{\kappa_g}$	ΔG_c from approx.[e] $\overline{K_m}, \overline{K_g}$	Approx. $\overline{K_m}$	Exact $\overline{\kappa_m}$	Approx. $\overline{K_g}$	Area
Methane	5.19	5.20	5.18	0.2905	0.2909	0.0846	148.5
Ethane	7.07	7.17	7.19	0.2647	0.2648	0.0685	181.5
Propane	9.14	8.99	9.11	0.2473	0.2476	0.0594	211.3
t n-Butane	10.96	10.81	10.99	0.2344	0.2345	0.0520	241.0
g n-Butane	10.67	10.69	10.93	0.2341	0.2348	0.0536	237.9
Isobutane	11.03	10.68	10.87	0.2344	0.2346	0.0536	237.1
t n-Pentane	13.12	12.65	12.90	0.2238	0.2241	0.0461	270.7
g n-Pentane	12.84	12.53	12.81	0.2237	0.2243	0.0472	267.8
Neopentane	12.69	12.13	12.44	0.2245	0.2253	0.0497	259.1
Cyclopentane	11.17	10.97	11.13	0.2318	0.2316	0.0531	239.4
Dimethylbutane	14.25	13.70	14.05	0.2159	0.2168	0.0453	282.5
Cyclohexane	12.55	12.32	12.53	0.2234	0.2230	0.0495	259.4
Dimethylpentane	16.01	15.67	16.11	0.2080	0.2092	0.0403	314.6
Cycloheptane	13.66	13.68	13.98	0.2154	0.2155	0.0464	279.1
Isooctane	17.35	16.96	17.52	0.2018	0.2032	0.0395	331.7
$CH_4 — CH_4$ 0.0[b]		5.20	5.18	0.2905	0.2909	0.0846	148.5
$CH_4 — CH_4$ 1.54[b]		7.18	7.21	0.2645	0.2648	0.0684	181.8
$CH_4 — CH_4$ 3.45[b]		9.39	9.62	0.2453	0.2469	0.0558	223.1
$CH_4 — CH_4$ 4.0[b]		9.91	10.30	0.2412	0.2445	0.0532	234.9
$CH_4 — CH_4$ 5.0[b]		10.62	11.54	0.2356	0.2438	0.0505	256.5
$CH_4 — CH_4$ 6.0[b]		10.73	12.68	0.2331	0.2507	0.0498	278.1
$CH_4 — CH_4$ 6.7[b]		9.59	13.37	0.2353	0.2687	0.0536	293.2
$CH_4 — CH_4$ 7.16[b]		7.72	12.80	0.2477	0.2909	0.0602	297.0
$CH_4 — CH_4$ 8.0[b]		7.72	11.36	0.2717	0.2909	0.0724	297.0
$CH_4 — CH_4$ 9.0[b]		7.72	10.50	0.2875	0.2909	0.0824	297.0
$CH_4 — CH_4$ 10.0[b]		7.72	10.35	0.2905	0.2909	0.0846	297.0

[a] Energies in kcal mol^{-1}, areas in Å^2, mean curvatures in Å^{-1}; Gaussian curvatures in Å^{-2}.
[b] Carbon–carbon internuclear distance in Å.
[c] This is ΔG_c from $\Delta G_c = \Delta G^* - e_i$.
[d] Total Gaussian curvature is 4π for all surfaces except last four. These have total Gaussian curvature 8π.
[e] Using the approximate mean curvature and the approximate Gaussian curvature of Appendix B with parameters -2.72 and 1.50.

It appears to be physically more appropriate to calculate an approximate curvature or a ''solvent accessibility of a solvent molecule'' in the manner of Nicholls et al.,[10] rather than an exact geometric curvature. Such an approximate treatment takes account of the proximity of another surface, whereas the more accurate geometrical treatment of Appendix A does not. An important biological example is an enzyme binding pocket, for example, chymotrypsin. When empty, the pocket may be at a higher energy than suggested by a purely geometric curvature.

In regard to the Gaussian curvature term in eq. (16), the surface curvature correction, $g(k)$, basically arises from the product $(1 - a/r_u)(1 - a/r_v)$, so that the Gaussian curvature $(1/r_u)(1/r_v)$ is an essential geometric component in the relationship between the areas A_a and A_m. In eq. (17), this term represents a contribution of $4\pi a^2\gamma_\infty$ for each topologically spherical surface. Empirically, it can be seen to be nonzero from the parameters necessary to fit the data of Table I, and in fact has a value of on the order of 2.5 kcal mol^{-1}.

If the Gaussian curvature term in eq. (16) is neglected, the data can still be fitted, but only if γ_∞ is considered an adjustable parameter.[24] The optimized values of 92 cal mol^{-1}Å^{-2} for this quantity and -2.13 for the remaining curvature coefficient gave a good fit of the data of Table I using the curvature $\overline{K_m}$. The problem with this, of course, is that the macroscopic limit is not the accepted value of 103.5 cal mol^{-1} Å^{-2} at 298 K.[25]

That the Gaussian curvature term should be included in the energy expression is further sub-

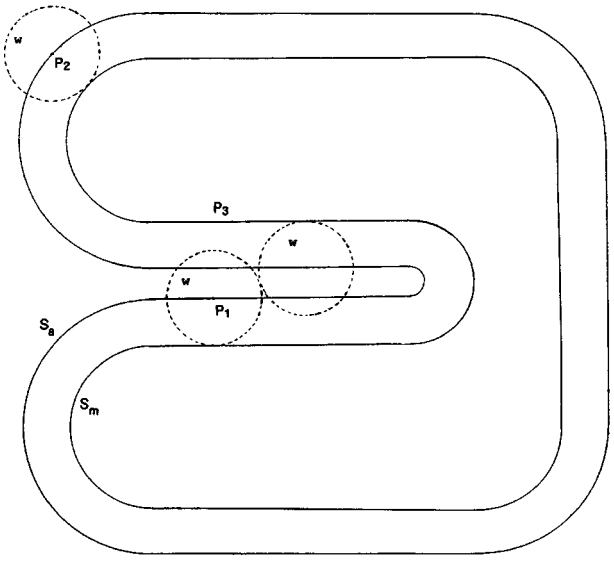

FIGURE 3. A molecular shape for which the use of the exact curvature will not give a good surface free energy.

stantiated by comparing eq. (16) with scaled particle theory. In the usual power series expansion,[12] in terms of the cavity radius, of G_c in scaled particle theory, the coefficient of the zeroth power (usually symbolized K_o) consists of two parts. The first is a term representing the free energy to introduce a point particle in solution, and neglecting the pressure volume term, there is a second term of about 1.3 kcal mol^{-1}. This second term corresponds to the Gaussian curvature term. If scaled particle theory is written in a form similar to eq. (16), that is, a surface energy times an area times an inverse radius power series, the coefficients of the inverse radius power series approach the coefficients of eq. (16) as the solvent sphere volume fraction $4\pi\rho a^2/3 \rightarrow 1$, if the point particle contribution and the pressure is neglected. ρ is the number density.

However, the use of the exact Gaussian curvature in the last term causes a problem in that the exact Gaussian curvature-produced contributions are discrete quantities, producing a jump discontinuity of $4\pi a^2\gamma_\infty$ when followed dynamically during a topological change of surface. A similar discontinuity problem arises for the molecular surface, as defined in Figure 1.

As an example of such a topological change of surface, eq. (16) does not have the correct behavior when the surface representing a dimer splits into two parts to form two monomers. There is a discontinuity in the Gaussian curvature of 4π when a second surface forms, giving rise to a discontinuous increase in ΔG_c. A further example of this discontinuity may be when normally toroidal HIV protease[26] (total Gaussian curvature = 0 for a torus) changes to a conformation where the active site flap is opened to produce a topologically spherical molecule.

To remedy this problem, the average Gaussian curvature was calculated approximately by the method of Appendix B. This approximate method, since it takes into account solvent accessibility of the solvent, rather than simply the curvature of the surface, is capable of giving a continuous range of values during topological changes in the surface. This behavior is different than using the exact Gaussian curvature, but agrees in the limiting case of distinct spherical molecules.

Using the approximate curvature, eq. (16) becomes

$$\Delta G_c = A_a\gamma_\infty\left(1 + b\overline{K_m} + c\overline{K_g}\right) \qquad (21)$$

where $\overline{K_m}$ is the average approximate mean curvature and $\overline{K_g}$ is the average approximate Gaussian curvature as found in Appendix B. Table I gives the results with optimized parameters $b = -2.72$ and $c = 1.50$. The parameter c also properly smoothes the discontinuous changes in the point particle contribution which are proportional to area times curvature.

Another equation that addresses all of the above problems which was used to calculate cavity formation energies successfully has been given as[27]

$$\Delta G_c = A_a\gamma_\infty\left(1 + b\overline{K_m} + c\overline{K_m^2}\right) \qquad (22)$$

where $\overline{K_m}$ is average approximate mean curvature, calculated in the manner of Nicholls et al.,[10] and $\overline{K_m^2}$ is the average of the square of the mean curvature. The mean curvature at a point was squared to get the squared mean curvature at that point. The parameters b and c were found to be -3.05 and 3.15, respectively, for an all atom treatment. This proved to be adequate in not only treating hydrocarbon solvation energies and hydrocarbon complexes but also in qualitatively following hydrophobic dissociation. Again, where two particles form where there was one, there is a smooth transition and the correct limiting energy (zero) is obtained in the limit of large separation.

Free energies of cavity formation, united atom cavity areas, and curvatures, calculated by two different methods, are given in Table I for selected

hydrocarbon systems. It can be seen from Figure 4 that the correlations are excellent for both methods but only slightly worse using the exact curvature. Such a series is easily correlated and thus not a good test of the model, which is designed to handle general surfaces.

The result of the two methods as applied to the methane–methane hydrophobic interaction is shown in Figure 5. This is a solvent effect only. The discontinuity can be seen in the exact curvature treatment. There is no discontinuity in the approximate curvature treatment. Using $\overline{K_m^2}$, which is easier to calculate than $\overline{K_g}$, gives a slightly higher barrier than using $\overline{K_g}$.[27]

Discussion

It has been shown that the areas A_m and A_a of parallel surfaces S_m and S_a are rigorously related if a particular curvature function is applied to A_a. Using the corresponding free energy expressions, the solubilities of small hydrocarbon molecules are easily correlated, but systems with complicated surfaces cannot be treated adequately unless an approximate rather than an exact curvature is used. The accessible surface area approach is expected to be better than the molecular surface area approach because the former contains a curvature expression capable of this necessary modification.

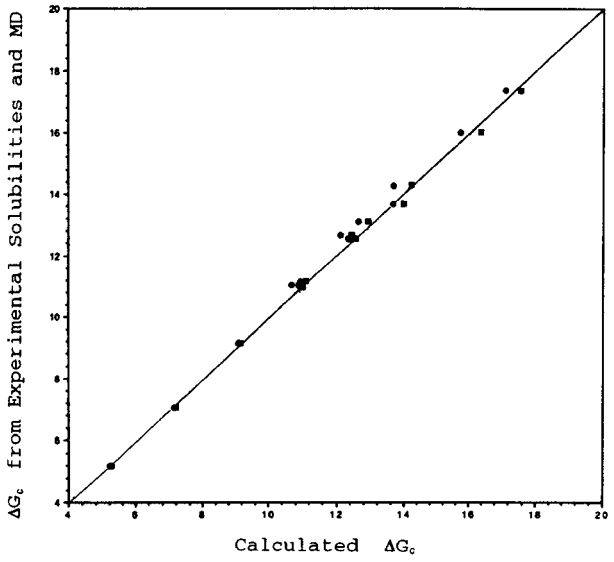

FIGURE 4. Calculated cavity potentials vs. cavity potentials from experimental solubilities and molecular dynamics. (●) eq. (16); (■) eq. (21).

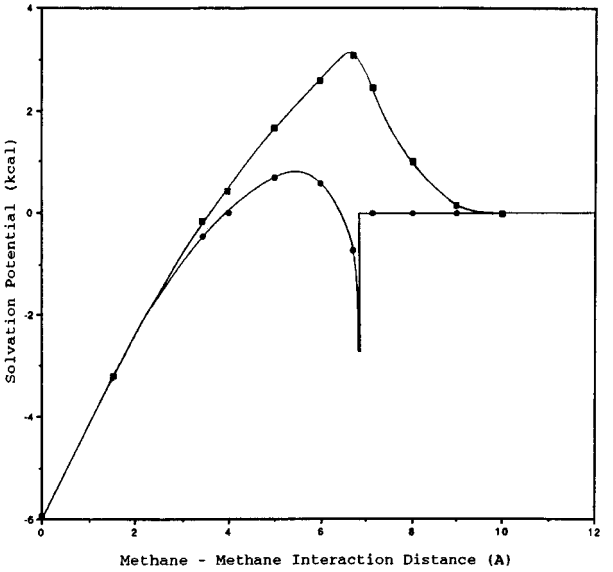

FIGURE 5. Methane–methane interaction. One accessible surface cavity is present up to 6.875 Å, then two cavities are formed. (●) eq. (16); (■) eq. (21). Solvation potential was calculated from $\Delta G^* = \Delta G_c + e_i$. Values of e_i from ref. 23. For 6.7 Å separation, e_i is interpolated from 6.0 Å and 7.16 Å values.

With this method, when two hydrocarbon molecules separate, the hydrophobic solvation energy passes over a barrier and goes smoothly to zero at large separation. The barrier of about 3 kcal mol^{-1} found here in the methane–methane dissociation curve is larger than that found from molecular dynamics calculations.[28–32] Part of the reason may be that a water molecule between the solutes can still interact strongly with the bulk solvent, while the model predicts that the relative inaccessibility of the spherical surface of that water molecule should result in raising the energy. While the curve representing the methane–methane dissociation using the exact curvature gives a lower barrier than that using approximate curvature, the use of the exact curvature is expected to fail in the general case. Eq. (14) indicates that this same failure is expected if the molecular surface is used.

The model outlined here should be an improvement over other attempts to parameterize scaled particle theory and apply it to solvation, because it takes into account molecular shapes and topological changes in the system, such as a change in the number of solutes. A number of difficulties remain. An implicit assumption is that the solvent density in the first solvation shell is uniform in the layer represented by or closest to the accessible

surface. It has been shown through molecular dynamics calculations that this density is not, in general, uniform.[27,29] Also, the parameters are considered to be temperature independent, so that the influence of temperature is through the surface tension parameter only. Using the accessible surface of a spherical solvent molecule to be simply proportional to its contribution to the surface free energy rather than including an angular dependence, is undoubtedly a poor approximation for water.

Appendix A

ALGORITHM FOR THE AVERAGE MEAN CURVATURE, $\overline{\kappa_m}$

To calculate the average of the mean curvature κ_m by a method that can be made accurate easily, it is necessary to calculate the mean curvature at each point or region on the accessible surface, and then average the area-weighted mean curvature for all regions of the whole surface.

The accessible surface is made up of a sum of atomic contributions A_{ai}. The contributions to the average mean curvature, $\overline{\kappa_{mi}}$, of each atom are weighted and averaged according to

$$\overline{\kappa_m} = (1/A_a) \sum_{i}^{\text{all atoms}} \overline{K_{mi}} A_{ai} \tag{23}$$

The curvature contribution of the accessible surface of an atom has two parts. The curvature of the spherical region is just the reciprocal of the spherical radius, but the boundary where the spherical region due to one atom meets the spherical region due to another atom also contributes to the curvature. The boundary between atoms represents a discontinuity in the tangent to the surface.

This accessible surface discontinuity gives a finite contribution to the mean curvature and requires a limiting approach treatment. The average mean curvature due to one atom is then

$$\overline{\kappa_{mi}} = (A_{ai}/r + D)/A_{ai} \tag{24}$$

where r is the sum of the radius of atom i and the solvent radius, and D is the contribution of the discontinuity at the boundary of the accessible area of atom i. The quantities r and D and other quantities in the remainder of this Appendix, are to be considered atomic rather than molecular quantities, that is, the subscript i will be understood rather than given explicitly.

The quantity D needs to be found by a limiting process and has the dimension area times curvature. It is the area of a portion of a torus times a curvature, after the limit is taken. Referring to Figure 6, the surface S'_a contains the torus region, and as the surface is allowed to approach the accessible surface S_a, the torus region area goes to zero, while its curvature goes to infinity. The area of a torus generated by rotating a circle of radius h, whose center is a distance R_p from the z axis, around the z axis as shown in Figure 6a, is given by[33,34]

$$A_t = \int_0^{2\pi}\int_0^{2\pi} h(R_p + h\cos\alpha)\, d\alpha\, d\theta \tag{25}$$

where the quantities appearing in the equation are defined in Figure 6a. For the full torus, the integration is over 2π for both angle parameters.

For the case of interest, as can be seen from Figure 6b, the integration over θ is from zero to 2π for a two-centered molecule, or, in the case of many centers, from zero to the angle representing the perimeter length of the discontinuity for each atom. The formula for the contribution of a single center to the discontinuity is then

$$D = \lim_{h\to 0} \int_0^{\theta_0}\int_{\pi}^{\phi+\pi/2} (1/2)(1/h + 1/R') \times h(R_p + h\cos\alpha)\, d\alpha\, d\theta \tag{26}$$

where the expression for the mean curvature [eq. (9)] is incorporated. The integration limits may be found from Figure 6b. R' is the radius of curvature which remains nonzero during the limiting process.

Letting h go to zero, and integrating

$$D = (\theta_0/2) R_p(\phi - \pi/2), \tag{27}$$

where, from Figure 6b, the intersection perimeter radius is

$$R_p = r\sin\phi \tag{28}$$

The following algorithm was used to get the angle θ_0 for each atom. To calculate the lengths of the spherical intersections of the spheres making up the molecule, each intersection circle was described by 5000 points, and each point was examined to see if it was inside or outside each neigh-

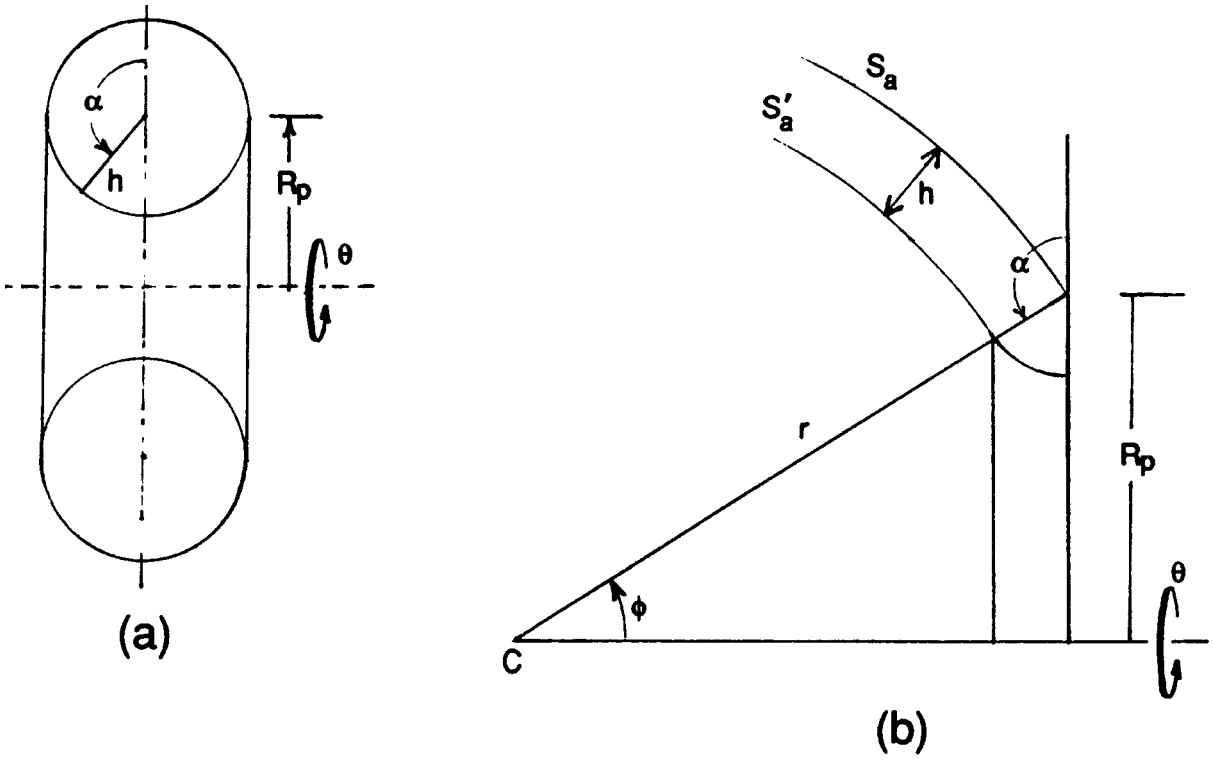

FIGURE 6. (a) Complete torus [see eq. (25)]. (b) Calculation of the curvature contribution of the discontinuity where two spherical parts of the accessible surface meet. Only one atom center is shown at C.

bor sphere. The number of points n gives the angle θ_0 by

$$\theta_0 = 2\pi n/5000. \tag{29}$$

When three atomic accessible surface sections intersect at a point, the point makes no contribution to the average mean curvature.

Appendix B

CALCULATION OF THE APPROXIMATE AVERAGE GAUSSIAN CURVATURE, $\overline{K_g}$

The calculation of the approximate Gaussian curvature, K_g, by a method analogous to the method for the approximate mean curvature[10] is done to avoid the jump discontinuity during topological changes and to be consistent with the real meaning of the approximate mean curvature, K_m, as the solvent accessibility of the solvent molecules.[10]

The curvature of the surface will in general vary at a point P depending on the direction. To calculate the Gaussian curvature at a point, it is necessary to find the two principal directions at that point. These are defined as the two directions at right angles in the surface which differ the most in curvature.

Analogous to the method of Nicholls, et al.,[10] for the mean curvature, an accessible surface is first obtained by placing a sphere of n points and of radius $r = R + a$ at each atomic center and removing all points closer to a center than the distance, here 3.4375 Å in the united atom treatment.

To find the Gaussian curvature of this surface at a given point P on the surface, a probe sphere of radius a_p and consisting of n points on its surface is centered at that point. For later computational convenience the molecule is oriented so that the point P and the atomic center to which it is related are placed on the z axis. The n points on the solvent sphere are then tested to see which ones fall inside or outside the molecular surface.

Referring to Figure 7, to find the two principal directions of curvature as mentioned previously, the points on the probe sphere are accordingly divided by arbitrarily m meridians into m groups or lunes with poles on the z axis. The curvature is then determined at the point P in $m/2$ directions in the surface, where each direction is defined by

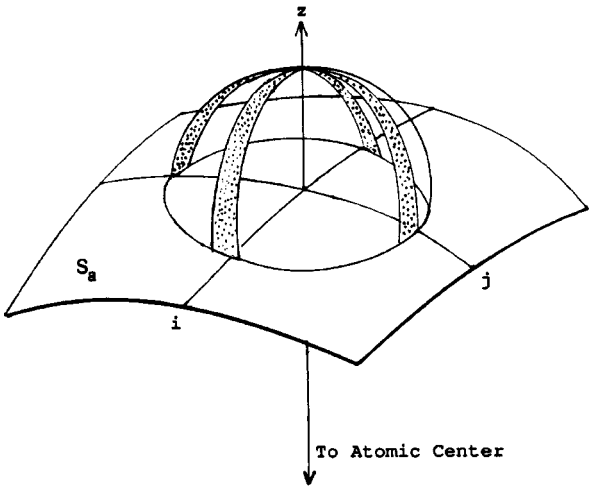

FIGURE 7. Calculation of the approximate Gaussian curvature. The methodology, consistent with that for the approximate mean curvature of Nichols *et al*. takes into account the possible proximity of other molecular regions. Four lunes are shown, two defining the curvature in direction i, and two defining the curvature in direction j.

two lunes 180° apart. For a probe sphere represented by n points, the curvature in one of the $m/2$ directions is

$$K_i = 2mV/na_p^2 - 2/a_p \tag{30}$$

where V is the number of exposed points (outside the accessible surface) in the two lunes 180° apart.

From this result, the curvature difference is found for each direction θ_i and the direction 90° to it; that is, the curvature in two directions 90° apart are subtracted giving $K_i - K_j$. The $m/4$ curvature differences are obtained in this manner. Of these $m/4$ differences, the largest absolute value is chosen, giving the two principal directions at the point P on the surface.

The Gaussian curvature at point P is then

$$K_g = K_i K_j \tag{31}$$

where i and j are the directions of maximum curvature difference (largest absolute $K_i - K_j$). The above procedure is repeated for each point P and averaged to get the average Gaussian curvature for the surface. Ideally, this average multiplied by the surface area will give 4π for a surface topologically equivalent to a sphere. Because this approximate method is dependent on the solvent accessibility of the solvent, it will give this only in limiting cases. For this work, n = 12,684, m = 24, and a_p = 2.8.

Acknowledgment

The author thanks Dr. Alex Varshavsky of Lilly Research Laboratories for a helpful discussion.

References

1. B. Lee and F. M. Richards, *J. Mol. Biol.*, **55**, 379 (1971).
2. F. M. Richards, *Annu. Rev. Biophys. Bioeng.*, **6**, 151 (1977).
3. I. Tunon, E. Silla, and J. L. Pascual-Ahuir, *Prot. Eng.*, **5**, 715 (1992).
4. R. M. Jackson and J. E. Sternberg, *Prot. Eng.*, **7**, 371 (1994).
5. J. L. Pascual-Ahuir, E. Sella, and J. Tunon, *J. Comput. Chem.*, **15**, 1127 (1994).
6. I. Langmuir, *Third Colloid Symposium Monograph*, Chemical Catalog Co., New York, 1925, p. 48.
7. S. C. Valvani, S. H. Yalkowski, and G. L. Amidon, *J. Phys. Chem.*, **80**, 829 (1976).
8. K. A. Sharp, A. Nicholls, R. F. Fine, and B. Honig, *Science*, **252**, 106 (1991).
9. R. C. Tolman, *J. Chem. Phys.*, **17**, 333 (1949).
10. A. Nicholls, K. A. Sharp, and B. Honig, *Prot. Struct. Funct. Genet.*, **11**, 281 (1991).
11. H. Reiss, H. L. Frisch, and J. L. Lebowitz, *J. Chem Phys.*, **31**, 369 (1959).
12. H. Reiss, *Adv. Chem. Phys.*, **9**, 1 (1965).
13. F. H. Stillinger, *J. Solution Chem.*, **2**, 141 (1973).
14. T. J. Willmore, *Differential Geometry*, Oxford University Press, Oxford, UK 1959, p. 116.
15. A. Goetz, *Introduction to Differential Geometry*, Addison-Wesley, Reading, MA, 1970, p. 268.
16. C. Tanford, *Proc. Nat. Acad. Sci. USA*, **76**, 4175 (1979).
17. R. D. Aveyard and A. Haydon, *J. Coll. Interf. Sci.*, **20**, 2255 (1965).
18. R. B. Hermann, *J. Phys. Chem.*, **76**, 2754 (1972).
19. R. B. Hermann, *Proc. Natl. Acad. Sci. USA*, **74**, 4144 (1977).
20. C. Chothia, *Nature*, **248**, 338 (1974).
21. A. Ben-Naim and Y. Marcus, *J. Chem. Phys.*, **81**, 2016 (1984).
22. G. Kaminski, E. M. Duffy, T. Matsui, and W. L. Jorgensen, *J. Phys. Chem.*, **98**, 13077 (1994).
23. R. B. Hermann, *J. Comput. Chem.*, **14**, 741 (1993).
24. Using γ_∞ as an adjustable parameter, ΔG^* itself rather than ΔG_c has been correlated with the accessible surface area without any curvature corrections, in D. J. Gieson, C. J. Cramer, and D. G. Truhlar, *J. Phys. Chem.*, **98**, 4141 (1994).
25. D. R. Lide, Ed., CRC Handbook of Chemistry and Physics, Chemical Rubber Publishing Co., Cleveland, OH, 1991.
26. Protein Data Bank, Brookhaven National Laboratory, Upton, NY.

27. R. B. Hermann, In *Structure and Reactivity in Aqueous Solution* (ACS Symposium Series No. 568), C. J. Cramer and D. G. Truhlar, Eds., 1994, p. 335.
28. W. L. Jorgensen, J. K. Buckner, S. Boudon, and J. Tirado-Rives, *J. Chem. Phys.*, **89**, 3742 (1988).
29. G. Ravishankar, M. Mezei, and D. L. Beverige, *Faraday Symp. Chem. Soc.*, **17**, 79 (1982).
30. D. E. Smith and A. D. J. Haymet, *J. Chem. Phys.*, **98**, 6445 (1993).
31. D. van Belle and S. J. Wodak, *J. Am. Chem. Soc.*, **115**, 647 (1993).
32. M. H. New and B. J. Berne, *J. Am. Chem. Soc.*, **117**, 7172 (1995).
33. E. Kreyszig, *Differential Geometry*, Dover, New York, 1991, p. 135.
34. M. P. do Carmo, *Differential Geometry ot Curves and Surfaces*, Prentice-Hall, Englewood Cliffs, NJ, 1976, p. 98.